

RECEIVED

MAY 07 2002

TECH CENTER 1600 2900



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

ppr 1-24

2 <110> APPLICANT: Corvas International , Inc.
 3 Vlasuk, George Phillip
 4 Stanssens, Patrick Eric Hugo
 5 Messens, Joris Hila Lieven
 6 Lauwereys, Marc Josef
 7 Laroche, Yves Rene
 8 Jespers, Laurent Stephane
 9 Gansemans, Yannick Georges Jozef
 10 Moyle, Matthew
 11 Bergum, Peter W.

*error on every
page. Please*

13 <120> TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE INHIBITORS AND

ANTICOAGULANT

14 PROTEIN
 16 <130> FILE REFERENCE: 018813/0272487
 18 <140> CURRENT APPLICATION NUMBER: 09/498,556A
 19 <141> CURRENT FILING DATE: 2002-04-10
 21 <150> PRIOR APPLICATION NUMBER: 08/809,455
 22 <151> PRIOR FILING DATE: 1997-04-17
 24 <150> PRIOR APPLICATION NUMBER: PCT/US95/13231
 25 <151> PRIOR FILING DATE: 1995-10-17
 27 <150> PRIOR APPLICATION NUMBER: 08/486,399
 28 <151> PRIOR FILING DATE: 1995-06-05
 30 <150> PRIOR APPLICATION NUMBER: 08/486,397
 31 <151> PRIOR FILING DATE: 1995-06-05
 33 <150> PRIOR APPLICATION NUMBER: 08/465,380
 34 <151> PRIOR FILING DATE: 1995-06-05
 36 <150> PRIOR APPLICATION NUMBER: 08/461,965
 37 <151> PRIOR FILING DATE: 1995-06-05
 39 <150> PRIOR APPLICATION NUMBER: 08/326,110
 40 <151> PRIOR FILING DATE: 1994-10-18
 42 <160> NUMBER OF SEQ ID NOS: 356
 44 <170> SOFTWARE: PatentIn version 3.1

consult

*Sequence
Rules for*

valid format

ERRORED SEQUENCES

61 <210> SEQ ID NO: 2
 62 <211> LENGTH: 228
 63 <212> TYPE: DNA
 W--> 64 <213> ORGANISM:
 66 <400> SEQUENCE: 2
 C--> 68 aaggcatacc cggagtgtgg tgagaatgaa tggctcgac tctgtggaac taagaagcca 60
 70 tgcgaggcca agtgcagtga ggaagaggag gaagatccga tatgccgac attttcttgt 120
 E--> 72 ccgggtcccc ctgcttgctg atgcgaagac ggattctaca gagacacggt gatcggcgac 180

→ This response is MANDATORY. This is a global

invalid

*error.
Edit all*

sequences.

See item 10

*on Error
summary sheet. 4/26/02*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

74 tgtgttaagg aagaagaatg cgaccaacat gagattatac atgtctga 228
 76 <210> SEQ ID NO: 3
 77 <211> LENGTH: (461) 460?
 78 <212> TYPE: DNA
 79 <213> ORGANISM: Ancylostoma caninum
 81 <220> FEATURE:
 82 <221> NAME/KEY: CDS
 83 <222> LOCATION: (22)..(321)
 85 <220> FEATURE:
 86 <221> NAME/KEY: misc_feature
 87 <223> OTHER INFORMATION:
 89 <400> SEQUENCE: 3

C--> 91 gaattccgct actactcaac a atg aag atg ctt tac gct atc gct ata atg 51
 92 Met Lys Met Leu Tyr Ala Ile Ala Ile Met
 W--> 93 15 10 misaligned nos. see item 3 on Ena Summary Sheet
 E--> 94 ttt ctc ctg gta tca tta tgc agc gca aga aca gtg agc aag gca tac 99
 95 Phe Leu Leu Val Ser Leu Cys Ser Ala Arg Thr Val Arg Lys Ala Tyr
 W--> 96 15 20 25 misaligned invalid
 97 ccg gag tgt ggt gag rat gaa tgg ctc gac gac tgt gga act gag aag 147
 W--> 98 Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Asp Cys Gly Thr Gin Lys
 W--> 99 30 35 40 misaligned invalid invalid
 E--> 100 cca tgc gag ccc aag tgc rat gag gra ccc cct gag gag gaa cat ccg (195)?
 101 Pro Cys Glu Ala Lys Cys Asn Glu Glu Pro Pro Glu Glu Glu Asp Pro
 W--> 102 45 50 55 misaligned invalid
 E--> 103 ata tgc cgc tca cgt ggt tgt tta tta cct cct gct tgc gta tgc aaa 243
 104 Ile Cys Arg Ser Arg Gly Cys Leu Leu Pro Pro Ala Cys Val Cys Lys
 W--> 105 60 65 70 misaligned nos. invalid
 E--> 106 gac cga ttc tac aga gac agc ctc atc ggc gac tct gtt agc gaa gaa 291
 W--> 107 Asp Gly Phe Tyr Arg Asp Trp Val Ile Oly Asp Cys Val Arg Glu Glu
 W--> 108 75 80 85 90 misaligned invalid invalid
 E--> 109 gaa tgc gac car cat gag att ata cat gtc t gaacgaqaaa gcaacaataa cc 344
 W--> 110 Glu Cys Asp Gin His Glu Ile Ile His Val invalid
 W--> 111 95 100 misaligned
 E--> 112 aaaggtcca actctcgtc tgcaaaatcg ctagtgtat gtctcttttg cgtccgaata 404
 E--> 114 gtttttagttg atgttaagta agaactctg ctggagagaa taaagcttc caactcc (461)?
 116 <210> SEQ ID NO: 4
 117 <211> LENGTH: 77
 118 <212> TYPE: PRT
 119 <213> ORGANISM: Ancylostoma caninum
 121 <400> SEQUENCE: 4
 123 Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Asp
 E--> 124 15 10
 125 Cys Gly Thr Gln Lys Pro Cys Glu Ala Lys Cys Asn Glu Glu
 E--> 126 15 20 25
 127 Pro Pro Glu Glu Glu Asp Pro Ile Cys Arg Ser Arg Gly Cys
 E--> 128 30 35 40
 129 Leu Leu Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg
 E--> 130 45 50 55
 131 Asp Thr Val Ile Gly Asp Cys Val Arg Glu Glu Glu Cys Asp

use lower-case letters
for all bases

must be
under
respective
codons

do not
use

tab codes

invalid

These
types

of errors
are global.

Please edit
globally.

misaligned

amino acid nos.

(see item 3
on Ena Summary
Sheet)

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

E--> 132 60 65 70 *misaligned amino acid nos.*

133 Gln His Glu Ile Ile His Val

E--> 134 75

214 <210> SEQ ID NO: 8

215 <211> LENGTH: 79

216 <212> TYPE: PRT

217 <213> ORGANISM: Ascylostoma caninum

219 <400> SEQUENCE: 8

221 Arg Thr Val Arg Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu

222 1 5 10 15

223 Asp Val Cys Gly Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu

224 20 25 30

225 Glu Glu Glu Asp Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala

226 35 40 45

227 Ala Cys Val Cys Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp

228 50 55 60

E--> 229 Cys Val Lys Glu Glu Glu Cys Asp Gln His Glu Ile (Ile) His (Val) *this is an upper-case I*

230 65 70 75 *invalid*

291 <210> SEQ ID NO: 10

292 <211> LENGTH: 425

293 <212> TYPE: DNA

294 <213> ORGANISM: Ancylostoma ceylanicum

296 <220> FEATURE:

297 <221> NAME/KEY: CDS (10) ← coding starts at location 10

298 <222> LOCATION: (1)..(291)

300 <220> FEATURE:

301 <221> NAME/KEY: misc_feature

302 <223> OTHER INFORMATION:

304 <400> SEQUENCE: 10

C--> 306 gaattccgc atg cgg acg ctc tac ctc att tct atc tgg ttg ttc ctc atc 51

OK 307 Met Arg Thr Leu Tyr Leu Ile Ser Ile Trp Leu Phe Leu Ile

308 1 5 10

309 tcg car tgt art gga aaa gga ttc ccg aaa tgt gac gtc art gaa aga 99

310 Ser Gln Cys Asn Gly Lys Ala Phe Pro Lys Cys Asp Val Asn Gln Arg

311 15 20 25 30

E--> 312 ttc ga0 gtg tgt ggc art ctg aag gag tgc gag ctc aag tgc gat gag 147 *invalid*

313 Phe Gln Val Cys Gly Asn Leu Lys Glu Cys Glu Leu Lys Cys Asp Glu

314 35 40 45

315 gac cct aag ata tgc tct cgt gca tgt att cgt ccc cct gct tgc gta 195

316 Asp Pro Lys Ile Cys Ser Arg Ala Cys Ile Arg Pro Pro Ala Cys Val

W--> 317 50 55 60 *misaligned nos. (see item 3)*

318 tgc gat gac gga ttc tac aga gac aaa tat ggc ttc tgt gtt gaa gaa 243

319 Cys Asp Asp Gly Phe Tyr Arg Asp Lys Tyr Gly Phe Cys Val Gln Gln

W--> 320 65 70 75 *misaligned nos.*

321 gac gaa tgt aac gat atg gag att att act ttt cca cca gaa acc aaa tg 293

322 Asp Glu Cys Asn Asp Met Glu Ile Ile Thr Phe Pro Pro Gln Thr Lys

W--> 323 80 85 90 *misaligned nos.*

325 atgaccgaag ctccacatt tctatacata tcttcactgc ttgacaggct tctcgacaat 353

327 ttagaagttc tgcttgactt tgtctatttg aaattgttca cactartggg ggaagtaaag 413

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

329 catttttcacg ac 425
 413 <210> SEQ ID NO: 13
 414 <211> LENGTH: 688
 415 <212> TYPE: DNA
 416 <213> ORGANISM: Ancylostoma ceylanicum
 418 <220> FEATURE:
 419 <221> NAME/KEY: CDS
 420 <222> LOCATION: (21)..(560)
 422 <220> FEATURE:
 423 <221> NAME/KEY: misc_feature
 424 <223> OTHER INFORMATION:
 426 <400> SEQUENCE: 13

C--> 429 gaattccggg cggcagaaag atg cga atg ctc ctt gtt cct atc tgg 50
 430 Met Arg Met Leu Tyr Leu Val Pro Ile Trp
 W--> 431 15 10 *modified nos.*
 432 ttg ctg ctc att tgg cta tgc agt gga aaa gct gcg aag aaa tgt ggt 98
 433 Leu Leu Leu Ile Ser Leu Cys Ser Gly Lys Ala Ala Lys Lys Cys Gly
 W--> 434 15 20 25
 435 ctc aat gaa agg ctg gac tgt ggc aat ctg aag caa tgc gag ccc aag 146
 436 Leu Asn Glu Arg Leu Asp Cys Gly Asn Leu Lys Gln Cys Glu Pro Lys
 W--> 437 30 35 40
 438 tgc agc gac ttg gaa agt gag gag tat gag gag gaa gat gag tgg aaa 194
 439 Cys Ser Asp Leu Glu Ser Glu Glu Tyr Glu Glu Glu Asp Glu Ser Lys
 W--> 440 45 50 55
 441 tgt cga tca cgt gaa tgt tct cgt cgt gtt tgt gta tgc gat gaa gga 242
 W--> 442 Cys Arg Ser Arg Glu Cys Ser Arg Arg Val Syc Val Syc Asp Glu Gly
 W--> 443 60 65 70 *initial*
 444 ttc tac aga aac aag aag ggc aag tgt gtt cga aaa gat gtt tgc gag 290
 445 Phe Tyr Arg Asn Lys Lys Gly Lys Cys Val Ala Lys Asp Val Cys Glu
 W--> 446 75 80 85 90
 447 gac gac aat atg gag att atc act ttt cca cca gaa gac gaa tgt ggt 338
 448 Asp Asp Asn Met Glu Ile Ile Thr Phe Pro Pro Glu Asp Glu Cys Gly
 W--> 449 95 100 105
 450 ccc gat gaa tgg ttc gac tac tgt gga aat tat aag aag tgc gaa cgc 386
 451 Pro Asp Glu Trp Phe Asp Tyr Cys Gly Asn Tyr Lys Lys Cys Glu Arg
 W--> 452 110 115 120
 453 aag tgc agt gag gag aca agt gag aaa aat gag gag gca tgc ctc tct 434
 454 Lys Cys Ser Glu Glu Thr Ser Glu Lys Asn Glu Glu Ala Cys Leu Ser
 W--> 455 125 130 135
 456 cgt gct tgt act ggt cgt gct tgc gta tgc aaa gac gga ttg tac aga 482
 457 Arg Ala Cys Thr Gly Arg Ala Cys Val Cys Lys Asp Gly Leu Tyr Arg
 W--> 458 140 145 150
 E--> 459 gac gac ttt ggc aac tgt gtt cca cat gac gaa tgc aac gat atg gag 530 *530/- insert*
 460 Asp Asp Phe Gly Asn Cys Val Pro His Asp Glu Cys Asn Asp Met Glu
 W--> 461 155 160 165 170
 462 atc atc act ttt cca ccg gaa acc aaa cat tgaccagagg ctccaactct cgct 584
 463 Ile Ile Thr Phe Pro Pro Glu Thr Lys His
 W--> 464 175 180
 465 acacaacgtc agggctagaa tggccctctt gcgagttagt agttttgctt gactctgctt 644

*use lower-case letters
for bases*

move under respective codons

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

467 atttgagcac ttctctattga tggcgaaaat aaagcattta aaac 688

469 <210> SEQ ID NO: 14

470 <211> LENGTH: (349) 301

471 <212> TYPE: DNA

472 <213> ORGANISM: Heligmosomoides polygyrus

474 <220> FEATURE:

475 <221> NAME/KEY: CDS

476 <222> LOCATION: (49)..(276)

478 <220> FEATURE:

479 <221> NAME/KEY: misc_feature

480 <223> OTHER INFORMATION:

482 <400> SEQUENCE: 14

C--> 484 gaattcgcgc cacctgagag gtgagctacg caagtcttcg ctggtaca atg atc cga 57

485 Met Ile Arg

486 1 move under codons

E--> 487 aag ctc gtt ctg act gct atc gtc acg gtg gtg cta agt gcg aag (105) 102

488 Lys Leu Val Leu Thr Ala Ile Val Thr Val Val Leu Ser Ala Lys

W--> 489 5 10 15 misaligned nos.

E--> 490 acc tgt gga cca aac gag gag tac act gaa tgc ggg acg cca tgc gag (153)

491 Thr Cys Gly Pro Asn Glu Glu Tyr Thr Glu Cys Gly Thr Pro Cys Glu

W--> 492 20 25 30 35 nos. off

E--> 493 ccg aag tgc aat gaa ccg atg cca gac atc tgt act ctg aac tgc atc (201)

494 Pro Lys Cys Asn Glu Pro Met Pro Asp Ile Cys Thr Leu Asn Cys Ile

W--> 495 40 45 50

E--> 496 gtg aac gtg tgt cag tgc aaa ccc ggc ttc aag cgc gga ccg aaa gga (249)

497 Val Asn Val Cys Gln Cys Lys Pro Gly Phe Lys Arg Gly Pro Lys Gly

W--> 498 55 60 65

E--> 499 tgc gtc gcc ccc gga cca ggc tgt aaa tagttctcca cctgcccttt cgttgga (304)

500 Cys Val Ala Pro Gly Pro Gly Cys Lys

E--> 501 70 75

588 <210> SEQ ID NO: 17

589 <211> LENGTH: (425) 434

590 <212> TYPE: DNA

591 <213> ORGANISM: Heligmosomoides polygyrus

593 <220> FEATURE:

594 <221> NAME/KEY: CDS (40) ← coding begins at location 40

595 <222> LOCATION: (1)..(291)

597 <220> FEATURE:

598 <221> NAME/KEY: misc_feature

599 <223> OTHER INFORMATION:

601 <400> SEQUENCE: 17

C--> 603 aagctttgct aacatactgc gtaataagga gtcttaatc atg cca gtt ctt ttg 54

604 Met Pro Val Leu Leu

605 1 5

606 ggt att ccg tta tta ttg cgt ttc ctg ggt ttc ctt ctg gta act ttg 102

607 Gly Ile Pro Leu Leu Leu Arg Phe Leu Gly Phe Leu Leu Val Thr Leu

608 10 15 20

609 ttc ggc tat ctg ctt act ttc ctt aaa aag ggc ttc ggt aag ata gct 150

W--> 610 Phe Gly Tyr Leu Leu Thr Phe Leu Lys Lys Gly Phe Gly Lys Ile Ala

invalid

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

```

611          25          30          35
612 att gct att tca ttg ttt ctt gct att att ggg ctt aac tca att 198
613 Ile Ala Ile Ser Leu Phe Leu Ala Leu Ile Ile Gly Leu Asn Ser Ile
614          40          45          50
615 ctt gtg ggt tat ctc tct gat att agc gca caa tta ccc tct gat ttt 246
616 Leu Val Gly Tyr Leu Ser Asp Ile Ser Ala Gln Leu Pro Ser Asp Phe
617          55          60          65
618 gtt cag ggc gtt cag tta att ctc ccg tct aat gcg ctt ccc tgt ttt 294
619 Val Gln Gly Val Gln Leu Ile Leu Pro Ser Asn Ala Leu Pro Cys Phe
620 70          75          80          85
621 tat gtt att ctc tct gta aag gct gct att ttc att ttt gac gtt aaa 342
622 Tyr Val Ile Leu Ser Val Lys Ala Ala Ile Phe Ile Phe Asp Val Lys
623          90          95          100
624 caa aaa atc gtt tct tat ttg gat tgg gat aaa ggt gga ggc tca ggc 390
625 Gln Lys Ile Val Ser Tyr Leu Asp Trp Asp Lys Gly Gly Ser Gly
626          105          110          115
627 gga tcggccaagt cggccatccc atatcacgcg gccgcggatc c 434
628 Gly
640 <210> SEQ ID NO: 19
641 <211> LENGTH: 425 430 (p.7)
642 <212> TYPE: DNA
643 <213> ORGANISM: Ancylostoma ceylanicum
645 <220> FEATURE:
646 <221> NAME/KEY: CDS
647 <222> LOCATION: (10)..(282)
649 <220> FEATURE:
W--> 650 <221> NAME/KEY: "W" stands for A or T This goes on 2237 line, since it
651 <223> OTHER INFORMATION: is lower-case letter for base
653 <400> SEQUENCE: 19
C--> 655 gaattccgg ctg gtw tcc tac tgc agt gga aaa gca acg atg cag tgt ggt 51
656          Leu Val Ser Tyr Cys Ser Gly Lys Ala Thr Met Gln Cys Gly
657          1          5          10
658 gag aat gaa aag tac gat tcg tgc ggt agc aag gag tgc gat aag aag 99
659 Glu Asn Glu Lys Tyr Asp Ser Cys Gly Ser Lys Glu Cys Asp Lys Lys
660 15          20          25          30
661 tgc aaa tat gac gga gtt gag gag gaa gac gac gag gaa cct aat gtg 147
662 Cys Lys Tyr Asp Gly Val Glu Glu Glu Asp Asp Glu Glu Pro Asn Val
663          35          40          45
664 cca tgc cta gta cgt gtg tgt cat caa gat tgc gta tgc gaa gaa gga 195
665 Pro Cys Leu Val Arg Val Cys His Gln Asp Cys Val Cys Glu Glu Gly
666          50          55          60
667 ttc tat aga aac aaa gat gac aaa tgt gta tca gca gaa gag tgc gaa 243
668 Phe Tyr Arg Asn Lys Asp Asp Lys Cys Val Ser Ala Glu Asp Cys Glu
669          65          70          75
670 ctt gac aat atg gac ttt ata tat ccc gga act cga aac tgaacgaagg ctc 295
671 Leu Asp Asn Met Asp Phe Ile Tyr Pro Gly Thr Arg Asn
672          80          85          90
673 cattcttgcg gcacaagatc gattgtctct cccctgcac tcagtagttt tgctacattg 355
675 tatatggtag caaaaaatta gcttagggag aataaaatct ttacctatat ttaatcaatg 415

```

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

E--> 677 aagtattctc tttct 430 ←

679 <210> SEQ ID NO: 20

680 <211> LENGTH: 100

681 <212> TYPE: PRT

682 <213> ORGANISM: Ancylostoma caninum

E--> 684 <400> SEQUENCE: 18 20 ← change to

686 Met Lys Met Leu Tyr Ala Ile Ala Ile Met Phe Leu Leu Val Ser Leu

687 1 5 10 15

688 Cys Ser Ala Arg Thr Val Arg Lys Ala Tyr Pro Glu Cys Gly Glu Asn

689 20 25 30

690 Glu Trp Leu Asp Asp Cys Gly Thr Gln Lys Pro Cys Glu Ala Lys Cys

691 35 40 45

692 Asn Glu Glu Pro Pro Glu Glu Glu Asp Pro Ile Cys Arg Ser Arg Gly

693 50 55 60

694 Cys Leu Leu Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg Asp

695 65 70 75 80

696 Thr Val Ile Gly Asp Cys Val Arg Glu Glu Glu Cys Asp Gln His Glu

697 85 90 95

698 Ile Ile His Val

699 100

E--> 762 <210> SEQ ID NO: 23

763 <211> LENGTH: 96

764 <212> TYPE: PRT

765 <213> ORGANISM: Ancylostoma ceylanicum

767 <400> SEQUENCE: 23

E--> 767 23

868 <210> SEQ ID NO: 29

869 <211> LENGTH: 76

870 <212> TYPE: PRT

871 <213> ORGANISM: Helogmosomoides polygyrus

873 <400> SEQUENCE: 29

875 Met Ile Arg Lys Leu Val Leu Leu Thr Ala Ile Val Thr Val Val Leu

876 1 5 10 15

877 Ser Ala Lys Thr Cys Gly Pro Asn Glu Glu Tyr Thr Glu Cys Gly Thr

878 20 25 30

E--> 879 Pro Cys Glu Pro Lys Cys Asn Glu Pro Met Pro Asp Ile Cys Thr (Len) 121

880 35 40 45

881 Asn Cys Ile Val Asn Val Cys Gln Cys Lys Pro Gly Phe Lys Arg Gly

882 50 55 60

883 Pro Lys Gly Cys Val Ala Pro Gly Pro Gly Cys Lys

884 65 70 75

1121 <210> SEQ ID NO: 36

1122 <211> LENGTH: 686 685 (p8)

1123 <212> TYPE: DNA

1124 <213> ORGANISM: Ancylostoma caninum

1126 <220> FEATURE:

1127 <221> NAME/KEY: CDS

1128 <222> LOCATION: (14)..(556)

1130 <220> FEATURE:

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

1131 <221> NAME/KEY: misc_feature
 1132 <223> OTHER INFORMATION:
 1134 <400> SEQUENCE: 36

E--> 1136 aattccgga aaa atg ctg atg ctc tac ctt gtt cct atc tgg ttg cta (49) 48
 1137 Met Leu Met Leu Tyr Leu Val Pro Ile Trp Leu Leu
 1138 1 5 10

E--> 1139 ctc att tcg caa tgc agt gga aaa tcc gcg aag aaa tgt ggt ctc aat (97) 48
 1140 Leu Ile Ser Gln Cys Ser Gly Lys Ser Ala Lys Lys Cys Gly Leu Asn
 1141 15 20 25 nos. off

E--> 1142 gaa aaa ttg gac tgt ggc aat ctg aag gca tgc gag aaa aag tgc agc (145)
 1143 Glu Lys Leu Asp Cys Gly Asn Leu Lys Ala Cys Glu Lys Lys Cys Ser
 1144 30 35 40

E--> 1145 gac ttg gac aat gag gag gat tat aag gag gaa gat gag tcg aaa tgc (193)
 1146 Asp Leu Asp Asn Glu Glu Asp Tyr Lys Glu Glu Asp Glu Ser Lys Cys
 1147 45 50 55 60

E--> 1148 cga tca cgt gaa tgt agt cgt cgt gtt tgt gta tgc gat gaa gga ttc (241)
 1149 Arg Ser Arg Glu Cys Ser Arg Arg Val Cys Val Cys Asp Glu Gly Phe

W--> 1150 65 70

E--> 1151 tac aga aac aag aag ggc caa tgt gtg aca aga gat gat tgc gag tat (289)
 1152 Tyr Arg Asn Lys Lys Gly Gln Cys Val Thr Arg Asp Asp Cys Glu Tyr
 1153 80 85 90

E--> 1154 gac aat atg gag att atc act ttt cca cca gaa gat aaa tgt ggt ccc (337)
 1155 Asp Asn Met Glu Ile Ile Thr Phe Pro Pro Glu Asp Lys Cys Gly Pro

W--> 1156 95 100 101

E--> 1157 gat gaa tgg ttc gac tgg tgt gga act tac aag cag tgt gag cgc aag (385)
 1158 Asp Glu Trp Phe Asp Trp Cys Gly Thr Tyr Lys Gln Cys Glu Arg Lys
 1159 110 115 120

E--> 1160 tgc aat aag gag cta agt gag aaa gat gaa gag gca tgc ctc tca cgt (433)
 1161 Cys Asn Lys Glu Leu Ser Glu Lys Asp Glu Glu Ala Cys Leu Ser Arg

W--> 1162 125 130 135

E--> 1163 gct tgt act ggt cgt gct tgt gtt tgc aac gac gga ctg tac aga gac (481)
 1164 Ala Cys Thr Gly Arg Ala Cys Val Cys Asn Asp Gly Leu Tyr Arg Asp
 1165 145 150 155

E--> 1166 gat ttt ggc aat tgt gtt gag aaa gac gaa tgt aac gat atg gag att (529)
 1167 Asp Phe Gly Asn Cys Val Glu Lys Asp Glu Cys Asn Asp Met Glu Ile
 1168 160 165 170

E--> 1169 atc act ttt cca ccg gaa acc aaa cac tgaccaaagg ctctaactct cgctacat (584)
 1170 Ile Thr Phe Pro Pro Glu Thr Lys His
 1171 175 180

E--> 1173 aacgtcagtg cttgaattgc cccctttacga gtttagtaatt ttgaetaact ctgtgtaatt (644)
 E--> 1175 gagcattgtc tactgatggt gaaaatgaag tgttcaatgt ct (686)

1178 <210> SEQ ID NO: 37
 1179 <211> LENGTH: 707
 1180 <212> TYPE: DNA
 1181 <213> ORGANISM: Ancylostoma caninum
 1183 <220> FEATURE:
 1184 <221> NAME/KEY: CDS
 1185 <222> LOCATION: (34)..(576) P. 9
 1187 <220> FEATURE:

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

use lower-case letters

1188 <221> NAME/KEY: misc_feature
 1189 <223> OTHER INFORMATION:
 1191 <400> SEQUENCE: 37

C--> 1193 **gaattccgcg gaattccggt tggcggcaga aaa atg ctg atg ctc tac ctt gtt** 54
 1194 Met Leu Met Leu Tyr Leu Val
 1195 1 5
 1196 cct atc tgg ttc ctg ctc att tcg caa tgc agt gga aaa tcc gcg aag 102
 1197 Pro Ile Trp Phe Leu Leu Ile Ser Glu Cys Ser Gly Lys Ser Ala Lys
 1198 10 15 20
 1199 aaa tgt ggc ctc aat gaa aaa ttg gac tgt ggc aat ctg aag gca tgc 150
 1200 Lys Cys Gly Leu Asn Glu Lys Leu Asp Cys Gly Asn Leu Lys Ala Cys
 1201 25 30 35
 1202 gag aaa aag tgc agc gac ttg gac aat gag gag gat tat ggg gag gaa 198
 1203 Glu Lys Lys Cys Ser Asp Leu Asp Asn Glu Glu Asp Tyr Gly Glu Glu
 1204 40 45 50 55
 1205 gat gag tcg aaa tgc cga tca cgt gaa tgt att ggt cgt gtt tgc gta 246
 1206 Asp Glu Ser Lys Cys Arg Ser Arg Glu Cys Ile Gly Arg Val Cys Val
 1207 60 65 70
 1208 tgc gat gaa gga ttc tac aga aac aag aag ggc caa tgt gtg aca aga 294
 1209 Cys Asp Glu Gly Phe Tyr Arg Asn Lys Lys Gly Gln Cys Val Thr Arg
 1210 75 80 85
 1211 gac gat tgc gag tat gac aat atg gag att atc act ttt cca cca gaa 342
 1212 Asp Asp Cys Glu Tyr Asp Asn Met Glu Ile Ile Thr Phe Pro Pro Glu
 1213 90 95 100
 1214 gat aaa tgt ggt ccc gat gaa tgg ttc gac tgg tgt gga act tac aag 390
 1215 Asp Lys Cys Gly Pro Asp Glu Trp Phe Asp Trp Cys Gly Thr Tyr Lys
 1216 105 110 115
 E--> 1217 **cag tgt gag cgc aag tgc agt gag gag cta** ~~aga~~ **gag aaa aat gag gag** 438
 1218 Gln Cys Glu Arg Lys Cys Ser Glu Glu Leu Ser Glu Lys Asn Glu Glu
 1219 120 125 130 135
 1220 gca tgc ctc tca cgt gct tgt act ggt cgt gct tgc gtt tgc aac gac 486
 1221 Ala Cys Leu Ser Arg Ala Cys Thr Gly Arg Ala Cys Val Cys Asn Asp
 1222 140 145 150
 1223 gga ttg tat aga gac gat ttt ggc aat tgt gtt gag aaa gac gaa tgt 534
 1224 Gly Leu Tyr Arg Asp Asp Phe Gly Asn Cys Val Glu Lys Asp Glu Cys
 1225 155 160 165
 1226 aac gat atg gag att atc act ttt cca ccg gaa acc aaa cac tgaccaaagg 586
 W--> 1227 **Asn Asp Met Glu Ile Ile Thr Phe Pro Pro Glu** **TSr** **Lys His**
 1228 170 175 180
 1230 ctctagctct cgctacataa cgctcagtgt tgaattgtcc ctttacgtgt tagtaatttt 646
 1232 gactaactct gtgtatttga gcattgtcta ctaatggtga aaatgaagct tttcaatgac 706
 1234 t 707
 1339 <210> SEQ ID NO: 41
 1340 <211> LENGTH: 75
 1341 <212> TYPE: PRT
 1342 <213> ORGANISM: Ancylostoma caninum
 1345 <220> FEATURE:
 1346 <221> NAME/KEY: misc_feature
 1347 <223> OTHER INFORMATION:

p10

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

1349 <400> SEQUENCE: 41
 1351 Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Val Cys Gly
 E--> 1352 1 5 10 15 *misaligned nos.*
 1353 Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu Glu Glu Glu Asp
 E--> 1354 20 25 30
 1355 Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala Ala Cys Val Cys
 E--> 1356 35 40 45
 1357 Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Lys Glu
 E--> 1358 50 55 60
 1359 Glu Glu Cys Asp Gln His Glu Ile Ile His Val
 E--> 1360 65 70 75
 1362 <210> SEQ ID NO: 42
 1363 <211> LENGTH: 74
 1364 <212> TYPE: PRT
 1365 <213> ORGANISM: Ancylostoma caninum
 1368 <220> FEATURE:
 1369 <221> NAME/KEY: misc_feature
 1370 <223> OTHER INFORMATION:
 1372 <400> SEQUENCE: 42
 1374 Arg Thr Ala Arg Lys Pro Pro Thr Cys Gly Glu Asn Glu Arg Val Glu
 E--> 1375 1 5 10 15 *misaligned*
 1376 Trp Cys Gly Lys Glu Cys Glu Ile Thr Cys Asp Asp Pro Asp Lys Ile
 E--> 1377 20 25 30
 1378 Cys Arg Ser Leu Ala Cys Pro Gly Pro Pro Ala Cys Val Cys Asp Asp
 E--> 1379 35 40 45
 1380 Gly Tyr Tyr Arg Asp Thr Asn Val Gly Leu Cys Val Gln Tyr Asp Glu
 E--> 1381 50 55 60
 1382 Cys Asn Asp Met Asp Ile Ile Met Val Ser
 E--> 1383 65 70
 1385 <210> SEQ ID NO: 43
 1386 <211> LENGTH: 88
 1387 <212> TYPE: PRT
 1388 <213> ORGANISM: Ancylostoma caninum
 1391 <220> FEATURE:
 1392 <221> NAME/KEY: misc_feature
 1393 <223> OTHER INFORMATION:
 1395 <400> SEQUENCE: 43
 1397 Lys Pro Ser Glu Lys Glu Cys Gly Pro His Glu Arg Leu Asp Cys Gly
 E--> 1398 1 5 10 15 *misaligned*
 1399 Asn Lys Lys Pro Cys Glu Arg Lys Cys Lys Ile Glu Thr Ser Glu Glu
 E--> 1400 20 25 30
 1401 Glu Asp Asp Tyr Glu Glu Gly Thr Glu Arg Phe Arg Cys Leu Leu Arg
 E--> 1402 35 40 45
 1403 Val Cys Asp Glu Pro Tyr Glu Cys Ile Cys Asp Asp Gly Tyr Tyr Arg
 E--> 1404 50 55 60
 1405 Asn Lys Lys Gly Glu Cys Val Thr Asp Asp Val Cys Glu Glu Asp Phe
 E--> 1406 65 70 75 80
 1407 Met Glu Phe Ile Thr Phe Ala Pro
 E--> 1408 85

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

```

1409 <210> SEQ ID NO: 44
1410 <211> LENGTH: 87
1411 <212> TYPE: PRT
1412 <213> ORGANISM: Ancylostoma caninum
1415 <220> FEATURE:
1416 <221> NAME/KEY: misc_feature
1417 <223> OTHER INFORMATION:
1419 <400> SEQUENCE: 44
1421 Arg Pro Glu Lys Lys Cys Gly Pro Gly Glu Arg Leu Ala Cys Gly Asn
E--> 1422 15 10 15 misaligned
1423 Lys Lys Pro Cys Glu Arg Lys Cys Lys Ile Glu Thr Ser Glu Glu Glu
E--> 1424 20 25 30
1425 Asp Asp Tyr Pro Glu Gly Thr Glu Arg Phe Arg Cys Leu Leu Arg Val
E--> 1426 35 40 45
1427 Cys Asp Glu Pro Tyr Glu Cys Ile Cys Asp Asp Gly Tyr Tyr Arg Asn
E--> 1428 50 55 60
1429 Lys Lys Gly Glu Cys Val Thr Asp Asp Val Cys Gln Glu Asp Phe Met
E--> 1430 65 70 75 80
1431 Glu Phe Ile Thr Phe Ala Pro
E--> 1432 85
1434 <210> SEQ ID NO: 45
1435 <211> LENGTH: 86
1436 <212> TYPE: PRT
1437 <213> ORGANISM: Ancylostoma caninum
1440 <220> FEATURE:
1441 <221> NAME/KEY: misc_feature
1442 <223> OTHER INFORMATION:
1444 <400> SEQUENCE: 45
1446 Arg Pro Glu Lys Lys Cys Gly Pro Gly Glu Arg Leu Asp Cys Ala Asn
E--> 1447 15 10 15 misaligned
1448 Lys Lys Pro Cys Glu Pro Lys Cys Lys Ile Glu Thr Ser Glu Glu Glu
E--> 1449 20 25 30
1450 Asp Asp Asp Val Glu Asp Thr Asp Val Arg Cys Leu Val Arg Val Cys
E--> 1451 35 40 45
1452 Glu Arg Pro Leu Lys Cys Ile Cys Lys Asp Gly Tyr Tyr Arg Asn Lys
E--> 1453 50 55 60
1454 Lys Gly Glu Cys Val Thr Asp Asp Val Cys Gln Glu Asp Phe Met Glu
E--> 1455 65 70 75 80
1456 Phe Ile Thr Phe Ala Pro
E--> 1457 85
1459 <210> SEQ ID NO: 46
1460 <211> LENGTH: 86
1461 <212> TYPE: PRT
1462 <213> ORGANISM: Ancylostoma caninum
1464 <220> FEATURE:
1465 <221> NAME/KEY: misc_feature
1466 <223> OTHER INFORMATION:
1468 <400> SEQUENCE: 46
1470 Arg Pro Glu Lys Lys Cys Gly Pro Gly Glu Arg Leu Asp Cys Ala Asn

```

p.12

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

E--> 1471 1 5 10 15 *misaligned nos.*
 1472 Lys Lys Pro Cys Glu Pro Lys Cys Lys Ile Glu Thr Ser Glu Glu Glu
 E--> 1473 20 25 30
 1474 Asp Asp Asp Val Glu Glu Thr Asp Val Arg Cys Leu Val Arg Val Cys
 E--> 1475 35 40 45
 1476 Glu Arg Pro Leu Lys Cys Ile Cys Lys Asp Gly Tyr Tyr Arg Asn Lys
 E--> 1477 50 55 60
 1478 Lys Gly Glu Cys Val Thr Asp Asp Val Cys Gln Glu Asp Phe Met Glu
 E--> 1479 65 70 75 80
 1480 Phe Ile Thr Phe Ala Pro
 E--> 1481 85
 1483 <210> SEQ ID NO: 47
 1484 <211> LENGTH: 78
 1485 <212> TYPE: PRT
 1486 <213> ORGANISM: Ancylostoma caninum
 1488 <220> FEATURE:
 1489 <221> NAME/KEY: misc_feature
 1490 <223> OTHER INFORMATION:
 1492 <400> SEQUENCE: 47
 1494 Lys Ser Leu Trp Asp Gln Lys Cys Gly Glu Asn Glu Arg Leu Asp Cys
 E--> 1495 1 5 10 15 *misaligned*
 1496 Gly Asn Gln Lys Asp Cys Glu Arg Lys Cys Asp Asp Lys Arg Ser Glu
 E--> 1497 20 25 30
 1498 Glu Glu Ile Met Gln Ala Cys Leu Thr Arg Gln Cys Leu Pro Pro Val
 E--> 1499 35 40 45
 1500 Cys Val Cys Glu Asp Gly Phe Tyr Arg Asn Asp Asn Asp Gln Cys Val
 E--> 1501 50 55 60
 1502 Asp Glu Glu Glu Cys Asn Met Glu Phe Ile Thr Phe Ala Pro
 E--> 1503 65 70 75
 1506 <210> SEQ ID NO: 48
 1507 <211> LENGTH: 89
 1508 <212> TYPE: PRT
 1509 <213> ORGANISM: Ancylostoma ceylanicum
 1511 <220> FEATURE:
 1512 <221> NAME/KEY: misc_feature
 1513 <223> OTHER INFORMATION:
 1515 <400> SEQUENCE: 48
 1517 Lys Pro Asn Asn Val Met Thr Asn Ala Cys Gly Leu Asn Glu Tyr Phe
 E--> 1518 1 5 10 15 *misaligned*
 1519 Ala Glu Cys Gly Asn Met Lys Glu Cys Glu His Arg Cys Asn Glu Glu
 E--> 1520 20 25 30
 1521 Glu Asn Glu Glu Arg Asp Glu Glu Arg Ile Thr Ala Cys Leu Ile Arg
 E--> 1522 35 40 45
 1523 Val Cys Phe Arg Pro Gly Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg
 E--> 1524 50 55 60
 1525 Asn Arg Thr Gly Ser Cys Val Glu Glu Asp Asp Cys Glu Tyr Glu Asn
 E--> 1526 65 70 75 80
 1527 Met Glu Phe Ile Thr Phe Ala Pro Glu
 E--> 1528 85

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

1530 <210> SEQ ID NO: 49
 1531 <211> LENGTH: 82
 1532 <212> TYPE: PRT
 1533 <213> ORGANISM: Ancylostoma ceylanicum
 1535 <220> FEATURE:
 1536 <221> NAME/KEY: misc_feature
 1537 <223> OTHER INFORMATION:
 1539 <400> SEQUENCE: 49
 1541 Val Pro Ile Cys Gly Ser Asn Glu Arg Tyr Ser Asp Cys Gly Asn Asp
 E--> 1542 1 5 10 15 *misaligned*
 1543 Lys Gln Cys Glu Arg Lys Cys Asn Glu Asp Asp Tyr Glu Lys Gly Asp
 E--> 1544 20 25 30
 1545 Glu Ala Cys Arg Ser His Val Cys Glu Arg Pro Gly Ala Cys Val Cys
 E--> 1546 35 40 45
 1547 Glu Asp Gly Phe Tyr Arg Asn Lys Lys Gly Ser Cys Val Glu Ser Asp
 E--> 1548 50 55 60
 1549 Asp Cys Glu Tyr Asp Asn Met Asp Phe Ile Thr Phe Ala Pro Glu Thr
 E--> 1550 65 70 75 80
 1551 Ser Arg
 1553 <210> SEQ ID NO: 50
 1554 <211> LENGTH: 84
 1555 <212> TYPE: PRT
 1556 <213> ORGANISM: Ancylostoma caninum
 1558 <220> FEATURE:
 1559 <221> NAME/KEY: misc_feature
 1560 <223> OTHER INFORMATION:
 1562 <400> SEQUENCE: 50
 1564 Lys Ser Ala Lys Lys Cys Gly Leu Asn Glu Lys Leu Asp Cys Gly Asn
 E--> 1565 1 5 10 15 *misaligned*
 1566 Leu Lys Ala Cys Glu Lys Lys Cys Ser Asp Leu Asp Asn Glu Glu Asp
 E--> 1567 20 25 30
 1568 Tyr Lys Glu Glu Asp Glu Ser Lys Cys Arg Ser Arg Glu Cys Ser Arg
 E--> 1569 35 40 45
 1570 Arg Val Cys Val Cys Asp Glu Gly Phe Tyr Arg Asn Lys Lys Gly Gln
 E--> 1571 50 55 60
 1572 Cys Val Thr Arg Asp Asp Cys Glu Tyr Asp Asn Met Glu Ile Ile Thr
 E--> 1573 65 70 75 80
 1574 Phe Pro Pro Glu
 1576 <210> SEQ ID NO: 51
 1577 <211> LENGTH: 84
 1578 <212> TYPE: PRT
 1579 <213> ORGANISM: Ancylostoma caninum
 1581 <220> FEATURE:
 1582 <221> NAME/KEY: misc_feature
 1583 <223> OTHER INFORMATION:
 1585 <400> SEQUENCE: 51
 1587 Lys Ser Ala Lys Lys Cys Gly Leu Asn Glu Lys Leu Asp Cys Gly Asn
 E--> 1588 1 5 10 15 *misaligned*
 1589 Leu Lys Ala Cys Glu Lys Lys Cys Ser Asp Leu Asp Asn Glu Glu Asp

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

E--> 1590 20 25 30 *misaligned*
 1591 Tyr Gly Glu Glu Asp Glu Ser Lys Cys Arg Ser Arg Glu Cys Ile Gly
 E--> 1592 35 40 45
 1593 Arg Val Cys Val Cys Asp Glu Gly Phe Tyr Arg Asn Lys Lys Gly Gln
 E--> 1594 50 55 60
 1595 Cys Val Thr Arg Asp Asp Cys Glu Tyr Asp Asn Met Glu Ile Ile Thr
 E--> 1596 65 70 75 80
 1597 Phe Pro Pro Glu
 1599 <210> SEQ ID NO: 52
 1600 <211> LENGTH: 83
 1601 <212> TYPE: PRT
 1602 <213> ORGANISM: Ancylostoma duodenale
 1604 <220> FEATURE:
 1605 <221> NAME/KEY: misc_feature
 1606 <223> OTHER INFORMATION:
 1608 <400> SEQUENCE: 52
 1610 Lys Ala Ala Lys Lys Cys Gly Leu Asn Glu Arg Leu Asp Cys Gly Asn
 E--> 1611 15 10 15 *misaligned*
 1612 Leu Lys Glu Cys Glu Pro Lys Cys Ser Asp Leu Glu Ser Glu Glu Tyr
 E--> 1613 20 25 30
 1614 Glu Glu Glu Asp Glu Ser Lys Cys Arg Ser Arg Glu Cys Ser Arg Arg
 E--> 1615 35 40 45
 1616 Val Cys Val Cys Asp Glu Gly Phe Tyr Arg Asn Lys Lys Gly Lys Cys
 E--> 1617 50 55 60
 1618 Val Ala Lys Asp Val Cys Glu Asp Asp Asn Met Glu Ile Ile Thr Phe
 E--> 1619 65 70 75 80
 1620 Pro Pro Glu
 1622 <210> SEQ ID NO: 53
 1623 <211> LENGTH: 78
 1624 <212> TYPE: PRT
 1625 <213> ORGANISM: Ancylostoma caninum
 1627 <220> FEATURE:
 1628 <221> NAME/KEY: misc_feature
 1629 <223> OTHER INFORMATION:
 1631 <400> SEQUENCE: 53
 1633 Asp Lys Cys Gly Pro Asp Glu Trp Phe Asp Trp Cys Gly Thr Tyr Lys
 E--> 1634 15 10 15 *misaligned*
 1635 Gln Cys Glu Arg Lys Cys Asn Lys Glu Leu Ser Glu Lys Asp Glu Glu
 E--> 1636 20 25 30
 1637 Ala Cys Leu Ser Arg Ala Cys Thr Gly Arg Ala Cys Val Cys Asn Asp
 E--> 1638 35 40 45
 1639 Gly Leu Tyr Arg Asp Asp Phe Gly Asn Cys Val Glu Lys Asp Glu Cys
 E--> 1640 50 55 60
 1641 Asn Asp Met Glu Ile Ile Thr Phe Pro Pro Glu Thr Lys His
 E--> 1642 65 70 75
 1644 <210> SEQ ID NO: 54
 1645 <211> LENGTH: 78
 1646 <212> TYPE: PRT
 1647 <213> ORGANISM: Ancylostoma caninum *P.15*

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

1649 <220> FEATURE:
 1650 <221> NAME/KEY: misc_feature
 1651 <223> OTHER INFORMATION:
 1653 <400> SEQUENCE: 54
 1655 Asp Lys Cys Gly Pro Asp Glu Trp Phe Asp Trp Cys Gly Thr Tyr Lys
 E--> 1656 15 10 15 *misaligned*
 1657 Gln Cys Glu Arg Lys Cys Ser Glu Glu Leu Ser Glu Lys Asn Glu Glu
 E--> 1658 20 25 30
 1659 Ala Cys Leu Ser Arg Ala Cys Thr Gly Arg Ala Cys Val Cys Asn Asp
 E--> 1660 35 40 45
 1661 Gly Leu Tyr Arg Asp Asp Phe Gly Asn Cys Val Glu Lys Asp Glu Cys
 E--> 1662 50 55 60
 1663 Asn Asp Met Glu Ile Ile Thr Phe Pro Pro Glu Thr Lys His
 E--> 1664 65 70 75
 1666 <210> SEQ ID NO: 55
 1667 <211> LENGTH: 77
 1668 <212> TYPE: PRT
 1669 <213> ORGANISM: Ancylostoma duodenale
 1671 <220> FEATURE:
 1672 <221> NAME/KEY: misc_feature
 1673 <223> OTHER INFORMATION:
 1675 <400> SEQUENCE: 55
 1677 Lys Cys Pro Thr Asp Glu Trp Phe Asp Trp Cys Gly Thr Tyr Lys His
 E--> 1678 15 10 15 *misaligned has.*
 1679 Cys Glu Leu Lys Cys Asp Arg Glu Leu Thr Glu Lys Glu Glu Gln Ala
 E--> 1680 20 25 30
 1681 Cys Leu Ser Arg Val Cys Glu Lys Ser Ala Cys Val Cys Asn Asp Gly
 E--> 1682 35 40 45
 1683 Leu Tyr Arg Asp Lys Phe Gly Asn Cys Val Glu Lys Asp Glu Cys Asn
 E--> 1684 50 55 60
 1685 Asp Met Glu Ile Ile Thr Phe Ala Pro Glu Glu Thr Lys
 E--> 1686 65 70 75
 1688 <210> SEQ ID NO: 56
 1689 <211> LENGTH: 78
 1690 <212> TYPE: PRT
 1691 <213> ORGANISM: Ancylostoma duodenale
 1694 <220> FEATURE:
 1695 <221> NAME/KEY: misc_feature
 1696 <223> OTHER INFORMATION:
 1698 <400> SEQUENCE: 56
 1700 Asp Glu Cys Gly Pro Asp Glu Trp Phe Asp Tyr Cys Gly Asn Tyr Lys
 E--> 1701 15 10 15 *misaligned has.*
 1702 Lys Cys Glu Arg Lys Cys Ser Glu Glu Thr Ser Glu Lys Asn Glu Glu
 E--> 1703 20 25 30
 1704 Ala Cys Leu Ser Arg Ala Cys Thr Gly Arg Ala Cys Val Cys Lys Asp
 E--> 1705 35 40 45
 1706 Gly Leu Tyr Arg Asp Asp Phe Gly Asn Cys Val Pro His Asp Glu Cys
 E--> 1707 50 55 60
 1708 Asn Asp Met Glu Ile Ile Thr Phe Pro Pro Glu Thr Lys His

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

E--> 1709 65 70 75
 1711 <210> SEQ ID NO: 57
 1712 <211> LENGTH: 75
 1713 <212> TYPE: PRT
 1714 <213> ORGANISM: Ancylostoma ceylanicum
 1717 <220> FEATURE:
 1718 <221> NAME/KEY: misc_feature
 1719 <223> OTHER INFORMATION:
 1721 <400> SEQUENCE: 57
 1723 Lys Ala Phe Pro Lys Cys Asp Val Asn Glu Arg Phe Glu Val Cys Gly
 E--> 1724 15 10 15 *misaligned*
 1725 Asn Leu Lys Glu Cys Glu Leu Lys Cys Asp Glu Asp Pro Lys Ile Cys
 E--> 1726 20 25 30
 1727 Ser Arg Ala Cys Ile Arg Pro Pro Ala Cys Val Cys Asp Asp Gly Phe
 E--> 1728 35 40 45
 1729 Tyr Arg Asp Lys Tyr Gly Phe Cys Val Glu Glu Asp Glu Cys Asn Asp
 E--> 1730 50 55 60
 1731 Met Glu Ile Ile Thr Phe Pro Pro Glu Thr Lys
 E--> 1732 65 70 75
 1734 <210> SEQ ID NO: 58
 1735 <211> LENGTH: 77
 1736 <212> TYPE: PRT
 1737 <213> ORGANISM: Ancylostoma ceylanicum
 1740 <220> FEATURE:
 1741 <221> NAME/KEY: misc_feature
 1742 <223> OTHER INFORMATION:
 1744 <400> SEQUENCE: 58
 1746 Arg Thr Val Lys Lys Cys Gly Lys Asn Glu Arg Tyr Asp Asp Cys Gly
 E--> 1747 15 10 15 *misaligned*
 1748 Asn Ala Tyr Asp Cys Glu Thr Lys Cys Gly Glu Glu Glu Lys Val Cys
 E--> 1749 20 25 30
 1750 Arg Ser Arg Glu Cys Thr Ser Pro Gly Ala Cys Val Cys Glu Gln Gly
 E--> 1751 35 40 45
 1752 Phe Tyr Arg Asp Pro Ala Gly Asp Cys Val Thr Asp Glu Glu Cys Asp
 E--> 1753 50 55 60
 1754 Glu Trp Asn Asn Met Glu Ile Ile Thr Met Pro Lys Gln
 E--> 1755 65 70 75
 1757 <210> SEQ ID NO: 59
 1758 <211> LENGTH: 84
 1759 <212> TYPE: PRT
 1760 <213> ORGANISM: Ancylostoma caninum
 1763 <220> FEATURE:
 1764 <221> NAME/KEY: misc_feature
 1765 <223> OTHER INFORMATION:
 1767 <400> SEQUENCE: 59
 1769 Lys Ala Thr Met Gln Cys Gly Glu Asn Glu Lys Tyr Asp Ser Cys Gly
 E--> 1770 15 10 15
 1771 Ser Lys Glu Cys Asp Lys Lys Cys Lys Tyr Asp Gly Val Glu Glu Glu
 E--> 1772 20 25 30

global env

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

1773 Asp Asp Glu Glu Pro Asn Val Pro Cys Leu Val Arg Val Cys His Glu

E--> 1774 (35 40 45)

1775 Asp Cys Val Cys Glu Glu Gly Phe Tyr Arg Asn Lys Asp Asp Lys Cys

E--> 1776 (50 55 60)

1777 Val Ser Ala Glu Asp Cys Glu Leu Asp Asn Met Asp Phe Ile Tyr Pro

E--> 1778 (65 70 75 80)

1779 Gly Thr Arg Asn

2064 <210> SEQ ID NO: 72

2065 <211> LENGTH: 5

2066 <212> TYPE: PRT

W--> 2067 <213> ORGANISM: ← mandatory response

2070 <220> FEATURE:

2071 <221> NAME/KEY: misc_feature

2072 <223> OTHER INFORMATION:

2074 <400> SEQUENCE: 72

E--> 2076 Gly (Trp Tyr Arg Asn) → invalid

E--> 2077 (15) ← mislabeled number

2154 <210> SEQ ID NO: 78

2155 <211> LENGTH: (5) 76

2156 <212> TYPE: PRT

W--> 2157 (213) ORGANISM: → response needed

2160 <220> FEATURE:

W--> 2161 <221> NAME/KEY: "Xaa" in locations 1 and 2 is an amino acid, provided that at

W--> 2162 least one Xaa is Glu or Asp

2163 <223> OTHER INFORMATION: ←

2165 <400> SEQUENCE: 78

W--> 2167 Xaa Xaa Gly Phe Tyr Arg Asp

E--> 2168 (15) ← mislabeled

2319 <210> SEQ ID NO: 89

2320 <211> LENGTH: (25) 22

2321 <212> TYPE: DNA

W--> 2322 (213) ORGANISM: → response needed

2324 <400> SEQUENCE: 89

E--> 2326 aaccataacc cccactcttc tc 22

2327 <210> SEQ ID NO: 90

2328 <211> LENGTH: 21

2329 <212> TYPE: PRT

W--> 2330 (213) ORGANISM: → response needed

2333 <220> FEATURE:

W--> 2334 <221> NAME/KEY: "Xaa" in locations 1 to 2 is an amino acid, move to <2237> line

2335 <223> OTHER INFORMATION: ←

2337 <400> SEQUENCE: 90

E--> 2339 AARCNTGYG ARMGGAARTG Y 21

2341 <210> SEQ ID NO: 91

2342 <211> LENGTH: 23

2343 <212> TYPE: PRT

2344 <213> ORGANISM: Ancylostoma caninum

2346 <220> FEATURE:

W--> 2347 <221> NAME/KEY: "W" stands for A or T; "R" stands for A of G; "N" stands for any ?

This is a peptide sequence.

It is already understood that Xaa is an amino acid.
any amino acid?
 (explain further)

Xaa can only represent an amino acid

There are bases, with cumulative base total.
(Per <2127>) This is a peptide sequence

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

W--> 2348 base; and "Y" stands for C or T. ?

2349 <223> OTHER INFORMATION:

2351 <400> SEQUENCE: 91

E--> 2351 TWRWANCCNT CYTTRCANAC RCA 23 ?

2354 <210> SEQ ID NO: 92

2355 <211> LENGTH: 13

2356 <212> TYPE: PRT

2357 <213> ORGANISM: Ancylostoma caninum

2359 <220> FEATURE:

2360 <221> NAME/KEY: misc_feature

2361 <223> OTHER INFORMATION:

2363 <400> SEQUENCE: 92

E--> 2365 Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Aop *invalid*

2366 1 5 10

2380 <210> SEQ ID NO: 94

2381 <211> LENGTH: 28 33

2382 <212> TYPE: DNA

2383 <213> ORGANISM: Ancylostoma caninum

2385 <220> FEATURE:

W--> 2386 <221> NAME/KEY: "R" stands for A or G; "N" stands for inosine; "Y" stands for C

W--> 2387 or T

2388 <223> OTHER INFORMATION: *goes on <2237> line*

2390 <400> SEQUENCE: 94

E--> 2392 aargentayc cngartgygg ngaraaygar tgg 33

2417 <210> SEQ ID NO: 97

2418 <211> LENGTH: 20

2419 <212> TYPE: PRT

2420 <213> ORGANISM: Ancylostoma caninum

2422 <220> FEATURE:

2423 <221> NAME/KEY: misc_feature

2424 <223> OTHER INFORMATION: N-terminal caninum

2426 <400> SEQUENCE: 97

E--> 2428 Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Tip *invalid* Leu Asp Asp Cys Gly Thr

2429 1 5 10 15

E--> 2430 Gb Lys Pro

2431 20

2529 <210> SEQ ID NO: 106

2530 <211> LENGTH: 36 5 shown below

2531 <212> TYPE: PRT

2532 <213> ORGANISM: Ancylostoma caninum

2534 <220> FEATURE:

2535 <221> NAME/KEY: misc_feature

2536 <223> OTHER INFORMATION: N-terminal fragment

2538 <400> SEQUENCE: 106

2540 Lys Ala Tyr Pro Glu

E--> 2541 1 5

2799 <210> SEQ ID NO: 128

2800 <211> LENGTH: 91

2801 <212> TYPE: PRT

p. 19

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

2802 <213> ORGANISM: Ancylostoma caninum
 2804 <220> FEATURE:
 2805 <221> NAME/KEY: misc_feature
 2806 <223> OTHER INFORMATION:
 2808 <400> SEQUENCE: 128
 E--> 2810 Leu Val Ear Tyr Cys Ser Gly Lys Ala Thr Met Gln Cys Gly Glu Asn
 2811 1 5 10 15
 2812 Glu Lys Tyr Asp Ser Cys Gly Ser Lys Glu Cys Asp Lys Lys Cys Lys
 2813 20 25 30
 2814 Tyr Asp Gly Val Glu Glu Glu Asp Asp Glu Glu Pro Asn Val Pro Cys
 2815 35 40 45
 2816 Leu Val Arg Val Cys His Gln Asp Cys Val Cys Glu Gly Phe Tyr
 2817 50 55 60
 2818 Arg Asn Lys Asp Asp Lys Cys Val Ser Ala Glu Asp Cys Glu Leu Asp
 2819 65 70 75 80
 2820 Asn Met Asp Phe Ile Tyr Pro Gly Thr Arg Asn
 2821 85 90
 2848 <210> SEQ ID NO: 131
 2849 <211> LENGTH: 6
 2850 <212> TYPE: DNA
 W--> 2851 <213> ORGANISM:
 2853 <220> FEATURE:
 W--> 2854 <221> NAME/KEY: Xaa in location 2 to 5 is an amino acid
 2855 <223> OTHER INFORMATION: Internal fragment
 2857 <400> SEQUENCE: 131
 W--> 2859 Cys Xaa Xaa Xaa Xaa Cys
 E--> 2860 1 5
 3425 <210> SEQ ID NO: 174
 3426 <211> LENGTH: 20
 3427 <212> TYPE: PRT
 W--> 3428 <213> ORGANISM:
 3430 <220> FEATURE:
 W--> 3431 <221> NAME/KEY: Xaa in locations 2 to 5 is an amino acid
 3432 <223> OTHER INFORMATION: Internal fragment
 3434 <400> SEQUENCE: 174
 W--> 3436 Cys Xaa Xaa Xaa Xaa
 E--> 3437 1 5
 3438 <210> SEQ ID NO: 175
 3439 <211> LENGTH: 20
 3440 <212> TYPE: PRT
 W--> 3441 <213> ORGANISM:
 3443 <220> FEATURE:
 W--> 3444 <221> NAME/KEY: Xaa in locations 2 to 4 is an amino acid
 3445 <223> OTHER INFORMATION: Internal fragment
 3447 <400> SEQUENCE: 175
 W--> 3449 Cys Xaa Xaa Xaa
 E--> 3450 1
 E--> 3490 <210> SEQ ID NO: 180
 3491 <211> LENGTH: 14

amino acids are properly listed

<212> response is DNA.

Why are amino acids shown below?

mandatory response

explain in another <223> line

explain further; Xaa can only be an amino acid, anyway.

explain further. show explanation on <223>

line, not on <221> line

Sequence 179 missing (p. 20)

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:42

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

3492 <212> TYPE: PRT
W--> 3493 <213> ORGANISM: *response needed*
3495 <220> FEATURE:
W--> 3496 <221> NAME/KEY: Xaa in locations 2 and locations 4 to 14 is an amino acid *what about 15?*
3497 <223> OTHER INFORMATION: Internal fragment
3499 <400> SEQUENCE: 180
W--> 3501 Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
E--> 3502 1 5 10 15
3503 <210> SEQ ID NO: 181
3504 <211> LENGTH: 8
3505 <212> TYPE: PRT
W--> 3506 <213> ORGANISM:
3508 <220> FEATURE:
W--> 3509 <221> NAME/KEY: Xaa in locations 2 to 8 is an amino acid
3510 <223> OTHER INFORMATION: Internal fragment
3512 <400> SEQUENCE: 181
W--> 3514 Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
E--> 3515 1 5 10
3516 <210> SEQ ID NO: 182
3517 <211> LENGTH: 7
3518 <212> TYPE: PRT
W--> 3519 <213> ORGANISM:
3521 <220> FEATURE:
W--> 3522 <221> NAME/KEY: Xaa in locations 2 to 7 is an amino acid
3523 <223> OTHER INFORMATION: Internal fragment
3525 <400> SEQUENCE: 182
W--> 3527 Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
E--> 3528 1 5
3529 <210> SEQ ID NO: 183
3530 <211> LENGTH: 6
3531 <212> TYPE: PRT
W--> 3532 <213> ORGANISM:
3534 <220> FEATURE:
W--> 3535 <221> NAME/KEY: Xaa in locations 2 to 6 is an amino acid
3536 <223> OTHER INFORMATION: Internal fragment
3538 <400> SEQUENCE: 183
W--> 3540 Cys Xaa Xaa Xaa Xaa Xaa Xaa
E--> 3541 1 5
3542 <210> SEQ ID NO: 184
3543 <211> LENGTH: 26
3544 <212> TYPE: PRT
W--> 3545 <213> ORGANISM:
3547 <220> FEATURE:
W--> 3548 <221> NAME/KEY: Xaa in locations 2 to 26 is an amino acid
3549 <223> OTHER INFORMATION: Internal fragment
3551 <400> SEQUENCE: 184
W--> 3553 Cys Xaa Xaa Xaa Xaa Xaa
E--> 3554 1 5
E--> 3832 <210> SEQ ID NO: 204

*Seq. 204 is shown twice
which one is
correct?*

*same
types of
error
in seqs 185-204
and in seqs 248-257*

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:43

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\L498556A.raw

E--> 3884 <210> SEQ ID NO: 207 *← shown here*
 E--> 3897 <210> SEQ ID NO: 209 *← seq. 208 missing*
 4220 <210> SEQ ID NO: 233
 4221 <211> LENGTH: 26
 4222 <212> TYPE: PRT
 W--> 4223 <213> ORGANISM:
 4225 <220> FEATURE:
 W--> 4226 <221> NAME/KEY: Xaa in location 2 to 26 is an amino acid *put in <223>*
 4227 <223> OTHER INFORMATION: Internal fragment
 4229 <400> SEQUENCE: 233
 W--> 4231 Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 4232 1 5 10 15
 E--> 4233 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 4234 20 25
 4235 <210> SEQ ID NO: 234
 4236 <211> LENGTH: 25
 4237 <212> TYPE: PRT
 W--> 4238 <213> ORGANISM:
 4240 <220> FEATURE:
 W--> 4241 <221> NAME/KEY: Xaa in location 2 to 25 is an amino acid *put in <223>*
 4242 <223> OTHER INFORMATION: Internal fragment
 4244 <400> SEQUENCE: 234
 W--> 4246 Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 4247 5 10 15
 E--> 4248 Xee Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 4249 20 25
 E--> 4454 <210> SEQ ID NO: 248 *shown twice*
 E--> 4561 <210> SEQ ID NO: 257 *seq. 256 missing*
 E--> 4717 <210> SEQ ID NO: 269 *seq. 268 missing*
 4821 <210> SEQ ID NO: 277
 4822 <211> LENGTH: 15
 4823 <212> TYPE: PRT
 W--> 4824 <213> ORGANISM:
 4826 <220> FEATURE:
 W--> 4827 <221> NAME/KEY: Xaa in location 2 and locations 4 to 15 is an amino acid *edit*
 4828 <223> OTHER INFORMATION: Internal fragment
 4830 <400> SEQUENCE: 277
 E--> 4832 Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 E--> 4833 1 5 10 15 *misaligned* *invalid* *delete*
 5349 <210> SEQ ID NO: 315
 5350 <211> LENGTH: 20
 5351 <212> TYPE: PRT
 W--> 5352 <213> ORGANISM:
 5354 <220> FEATURE:
 W--> 5355 <221> NAME/KEY: Xaa in locations 1 to 3 and locations 5 to 20 is an amino acid *edit*
 5356 <223> OTHER INFORMATION: Internal fragment
 5358 <400> SEQUENCE: 315
 E--> 5360 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 E--> 5361 1 5 10 15 *invalid*

Misaligned nos.

RAW SEQUENCE LISTING

DATE: 04/26/2002

PATENT APPLICATION: US/09/498,556A

TIME: 11:49:43

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

W--> 5362 Xaa Xaa Xaa Xaa
E--> 5363 (20) *misalign*
5904 <210> SEQ ID NO: 356
5905 <211> LENGTH: 6
5906 <212> TYPE: PRT
W--> 5907 <213> ORGANISM:
5909 <220> FEATURE:
W--> 5910 <221> NAME/KEY: Xaa in location 2 to 6 is an amino acid *edit*
5911 <223> OTHER INFORMATION: Internal fragment
5913 <400> SEQUENCE: 356
W--> 5915 Cys Xaa Xaa Xaa Xaa Xaa
E--> 5916 1 (5) *misalign*
E--> 5920 (. continued)
E--> 5921 (. continued)
E--> 5922 C:\WINNT\Profiles\chipman_1\Desktop\70016337_1.DOC

delete.

*See next page for more
error*

01/498,556A 23

<210> 255
<211> 7 6 shown
<212> PRT
<213>

<220>
<221> Xaa in location 2 to 6 is an amino acid
<223> Internal fragment

<400> 255

Cys Xaa Xaa Xaa Xaa Cys
1 5

Cys Xaa Xaa Xaa Cys
1

what is this?

5

FyI

1. The first step is to identify the amino acid sequence of the protein. This is done by comparing the protein sequence to a database of known sequences. 2. The second step is to identify the internal fragment. This is done by comparing the protein sequence to a database of known internal fragments. 3. The third step is to identify the amino acid sequence of the internal fragment. This is done by comparing the internal fragment sequence to a database of known sequences.

24

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002
TIME: 11:49:44

Input Set : A:\PTO.VSK.txt
Output Set: N:\CRF3\04262002\I498556A.raw

Skipped Sequences(NEW RULES):

Sequence(s) __missing. If intentional, please use the following format for each skipped sequence.

<210> sequence id number

<400> sequence id number

000

Seq#: 179, 208, 256, 268

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:19 M:271 C: Current Filing Date differs, Replaced Current Filing Date
 L:49 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:53 M:112 C: (48) String data converted to lower case,
 M:112 Repeated in SeqNo=1
 L:64 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:68 M:112 C: (48) String data converted to lower case,
 M:112 Repeated in SeqNo=2
 L:72 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 L:91 M:112 C: (48) String data converted to lower case,
 L:93 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
 L:94 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 M:112 Repeated in SeqNo=3
 L:96 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
 L:98 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:2
 L:99 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
 L:100 M:254 E: No. of Bases conflict, LENGTH:Input:195 Counted:194 SEQ:3
 L:100 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:2
 L:102 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
 M:254 Repeated in SeqNo=3
 L:103 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 L:105 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
 L:106 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:5
 L:107 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:3
 L:108 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
 L:109 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 L:110 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:2
 L:111 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
 L:112 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 L:114 M:252 E: No. of Seq. differs, <211> LENGTH:Input:461 Found:460 SEQ:3
 L:124 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:4
 M:332 Repeated in SeqNo=4
 L:151 M:112 C: (48) String data converted to lower case,
 M:112 Repeated in SeqNo=5
 L:170 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1
 L:229 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:247 M:112 C: (48) String data converted to lower case,
 M:112 Repeated in SeqNo=9
 L:260 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1
 L:264 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:267 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:270 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:273 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:276 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:279 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:282 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:285 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
 L:306 M:112 C: (48) String data converted to lower case,
 L:307 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 10, CDS LOCATION:1..291

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

M:112 Repeated in SeqNo=10

L:312 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1

L:317 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:10

L:320 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:10

L:323 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:10

L:347 M:112 C: (48) String data converted to lower case,

L:348 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1

L:349 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

M:112 Repeated in SeqNo=11

L:352 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

L:355 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

L:358 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

L:361 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

L:364 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

L:367 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

L:390 M:112 C: (48) String data converted to lower case,

L:393 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:12

M:112 Repeated in SeqNo=12

L:396 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:12

L:399 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:12

L:402 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:12

L:405 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:12

L:429 M:112 C: (48) String data converted to lower case,

L:431 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

M:112 Repeated in SeqNo=13

L:434 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:437 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:440 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:442 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:2

L:443 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:446 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:449 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:452 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:455 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:458 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:459 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:530 SEQ:13

L:461 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:464 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:13

L:484 M:112 C: (48) String data converted to lower case,

L:487 M:254 E: No. of Bases conflict, LENGTH:Input:105 Counted:102 SEQ:14

M:112 Repeated in SeqNo=14

L:489 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:14

M:254 Repeated in SeqNo=14

L:492 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:14

L:495 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:14

L:498 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:14

L:501 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:14

L:501 M:252 E: No. of Seq. differs, <211> LENGTH:Input:349 Found:301 SEQ:14

L:519 M:112 C: (48) String data converted to lower case,

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:520 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1
M:112 Repeated in SeqNo=15
L:561 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=16
L:603 M:112 C: (48) String data converted to lower case,
L:604 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 17, CDS LOCATION:1..291
M:112 Repeated in SeqNo=17
L:610 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1
L:628 M:252 E: No. of Seq. differs, <211> LENGTH:Input:425 Found:434 SEQ:17
L:633 M:201 W: Mandatory field data missing, <213> ORGANISM
L:650 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:19
L:655 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=19
L:677 M:252 E: No. of Seq. differs, <211> LENGTH:Input:425 Found:430 SEQ:19
L:684 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:20 differs:18
L:762 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO:23
L:767 M:252 E: No. of Seq. differs, <211> LENGTH:Input:96 Found:0 SEQ:23
L:879 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
L:889 M:201 W: Mandatory field data missing, <213> ORGANISM
L:893 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=30
L:917 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=31
L:961 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=32
L:1005 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=33
L:1048 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=34
L:1092 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=35
L:1136 M:254 E: No. of Bases conflict, LENGTH:Input:49 Counted:48 SEQ:36
M:112 Repeated in SeqNo=36
M:254 Repeated in SeqNo=36
L:1150 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:1156 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:1162 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:36
L:1175 M:252 E: No. of Seq. differs, <211> LENGTH:Input:686 Found:685 SEQ:36
L:1193 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=37
L:1217 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
L:1227 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1
L:1251 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=38
L:1295 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=39
L:1352 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:41
M:332 Repeated in SeqNo=41
L:1375 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:42

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

M:332 Repeated in SeqNo=42
 L:1398 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:43
 M:332 Repeated in SeqNo=43
 L:1422 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:44
 M:332 Repeated in SeqNo=44
 L:1447 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:45
 M:332 Repeated in SeqNo=45
 L:1471 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:46
 M:332 Repeated in SeqNo=46
 L:1495 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:47
 M:332 Repeated in SeqNo=47
 L:1518 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:48
 M:332 Repeated in SeqNo=48
 L:1542 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:49
 M:332 Repeated in SeqNo=49
 L:1565 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:50
 M:332 Repeated in SeqNo=50
 L:1588 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:51
 M:332 Repeated in SeqNo=51
 L:1611 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:52
 M:332 Repeated in SeqNo=52
 L:1634 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:53
 M:332 Repeated in SeqNo=53
 L:1656 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:54
 M:332 Repeated in SeqNo=54
 L:1678 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:55
 M:332 Repeated in SeqNo=55
 L:1701 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:56
 M:332 Repeated in SeqNo=56
 L:1724 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:57
 M:332 Repeated in SeqNo=57
 L:1747 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:58
 M:332 Repeated in SeqNo=58
 L:1770 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:59
 M:332 Repeated in SeqNo=59
 L:1793 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:60
 M:332 Repeated in SeqNo=60
 L:1814 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:61
 M:332 Repeated in SeqNo=61
 L:1835 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:62
 M:332 Repeated in SeqNo=62
 L:1870 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:63
 M:332 Repeated in SeqNo=63
 L:1904 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:64
 M:332 Repeated in SeqNo=64
 L:1962 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:1969 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:66
 L:1975 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66 after pos.:0
 L:1982 M:201 W: Mandatory field data missing, <213> ORGANISM

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:1989 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:67
 L:1994 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67 after pos.:0
 L:2000 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2007 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:68
 L:2008 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:68
 L:2009 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:68
 L:2014 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68 after pos.:0
 L:2020 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2035 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2050 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2067 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2076 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:2082 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2098 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2113 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2127 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2142 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2157 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2161 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:78
 L:2162 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:78
 L:2167 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:78
 L:2167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78 after pos.:0
 L:2168 M:252 E: No. of Seq. differs, <211> LENGTH:Input:5 Found:7 SEQ:78
 L:2173 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2177 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:79
 L:2178 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:79
 L:2183 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:79
 L:2183 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79 after pos.:0
 L:2189 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2193 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:80
 L:2194 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:80
 L:2199 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:80
 L:2199 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80 after pos.:0
 L:2205 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2209 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:81
 L:2210 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:81
 L:2215 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:81
 L:2215 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:81 after pos.:0
 L:2221 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2225 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:82
 L:2226 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:82
 L:2231 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:82
 L:2231 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82 after pos.:0
 L:2237 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2241 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:83
 L:2246 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:83
 L:2246 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83 after pos.:0
 L:2252 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2256 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:84

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:2257 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:84
 L:2258 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:84
 L:2263 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:84
 L:2263 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84 after pos.:0
 L:2269 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2273 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:85
 L:2278 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:85
 L:2278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85 after pos.:0
 L:2284 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2288 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:86
 L:2293 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:86
 L:2293 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86 after pos.:0
 L:2299 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2303 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:87
 L:2308 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:87
 L:2308 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87 after pos.:0
 L:2314 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2318 M:112 C: (48) String data converted to lower case.
 L:2322 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2326 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 L:2326 M:112 C: (48) String data converted to lower case.
 L:2326 M:252 E: No. of Seq. differs, <211> LENGTH:Input:25 Found:22 SEQ:89
 L:2330 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2334 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:90
 L:2339 M:252 E: No. of Seq. differs, <211> LENGTH:Input:21 Found:0 SEQ:90
 L:2347 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:91
 L:2348 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:91
 L:2353 M:252 E: No. of Seq. differs, <211> LENGTH:Input:23 Found:0 SEQ:91
 L:2365 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:2386 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:94
 L:2387 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:94
 L:2392 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:94
 L:2392 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:94 after pos.:0
 L:2392 M:112 C: (48) String data converted to lower case.
 L:2392 M:252 E: No. of Seq. differs, <211> LENGTH:Input:28 Found:33 SEQ:94
 L:2408 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2428 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:2430 M:333 E: Wrong sequence grouping, Amino acids not in groups!
 L:2430 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:2541 M:252 E: No. of Seq. differs, <211> LENGTH:Input:36 Found:5 SEQ:106
 L:2607 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2618 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2630 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2642 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2654 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2666 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2678 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2681 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:118
 L:2686 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:118

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:2686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118 after pos.:0
 L:2691 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2694 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:119
 L:2695 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:119
 L:2696 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:119
 L:2701 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:119
 L:2701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:119 after pos.:0
 L:2706 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2718 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2730 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2742 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2754 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2766 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2774 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 L:2778 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2790 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2798 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
 L:2810 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:2825 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2828 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:129
 L:2833 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:129
 L:2833 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:129 after pos.:0
 L:2838 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2846 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:130
 L:2846 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:130 after pos.:0
 L:2851 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2854 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:131
 L:2859 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:131
 L:2859 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:131 after pos.:0
 L:2860 M:252 E: No. of Seq. differs, <211> LENGTH:Input:6 Found:0 SEQ:131
 L:2864 M:201 W: Mandatory field data missing, <213> ORGANISM
 L:2867 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:132
 L:2872 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:132
 L:2872 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:132 after pos.:0
 L:2880 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:133
 L:2885 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:133
 L:2885 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:133 after pos.:0
 L:2893 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:134
 L:2898 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:134
 L:2898 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:134 after pos.:0
 L:2900 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:134
 L:2900 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:134 after pos.:16
 L:2908 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:135
 L:2913 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:135
 L:2913 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:135 after pos.:0
 L:2915 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:135
 L:2915 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:135 after pos.:16
 L:2923 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:136
 L:2928 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:136

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:2928 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:136 after pos.:0
 L:2930 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:136
 L:2930 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:136 after pos.:16
 L:2937 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:137
 L:2942 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:137
 L:2942 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:137 after pos.:0
 L:2944 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:137
 L:2944 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:137 after pos.:16
 L:2951 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:138
 L:2956 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:138
 L:2956 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:138 after pos.:0
 L:2958 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:138
 L:2958 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:138 after pos.:16
 L:2965 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:139
 L:2970 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:139
 L:2970 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139 after pos.:0
 L:2978 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:140
 L:2983 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:140
 L:2983 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140 after pos.:0
 L:2991 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:141
 L:2996 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:141
 L:2996 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:141 after pos.:0
 L:3004 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:142
 L:3009 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:142
 L:3009 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:142 after pos.:0
 L:3017 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:143
 L:3022 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:143
 L:3022 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:143 after pos.:0
 L:3030 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:144
 L:3035 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:144
 L:3035 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:144 after pos.:0
 L:3043 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:145
 L:3048 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:145
 L:3048 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:145 after pos.:0
 L:3056 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:146
 L:3061 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:146
 L:3061 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:146 after pos.:0
 L:3069 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:147
 L:3074 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:147
 L:3074 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:147 after pos.:0
 L:3087 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:148
 L:3087 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:148 after pos.:0
 L:3100 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:149
 L:3100 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:149 after pos.:0
 L:3113 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:150
 L:3113 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:150 after pos.:0
 L:3126 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:151
 L:3126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:151 after pos.:0
 L:3139 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:152

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002
TIME: 11:49:44

RECEIVED

MAY 9 8 2002

TECH CENTER 1600-2900

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:3139 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:152 after pos.:0
L:3152 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:153
L:3152 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:153 after pos.:0
L:3165 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:154
L:3165 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:154 after pos.:0
L:3178 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:155
L:3178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155 after pos.:0
L:3191 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:156
L:3191 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156 after pos.:0
L:3204 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:157
L:3204 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157 after pos.:0
L:3217 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:158
L:3230 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:159
L:3243 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:160
L:3437 M:252 E: No. of Seq. differs, <211> LENGTH:Input:20 Found:5 SEQ:174
L:3450 M:252 E: No. of Seq. differs, <211> LENGTH:Input:20 Found:4 SEQ:175
L:3490 M:214 E: (33) Seq.# missing, SEQ ID NO:179
L:3502 M:252 E: No. of Seq. differs, <211> LENGTH:Input:14 Found:15 SEQ:180
L:3515 M:252 E: No. of Seq. differs, <211> LENGTH:Input:8 Found:14 SEQ:181
L:3528 M:252 E: No. of Seq. differs, <211> LENGTH:Input:7 Found:8 SEQ:182
L:3541 M:252 E: No. of Seq. differs, <211> LENGTH:Input:6 Found:7 SEQ:183
L:3554 M:252 E: No. of Seq. differs, <211> LENGTH:Input:26 Found:6 SEQ:184
L:3569 M:252 E: No. of Seq. differs, <211> LENGTH:Input:25 Found:26 SEQ:185
L:3584 M:252 E: No. of Seq. differs, <211> LENGTH:Input:24 Found:25 SEQ:186
L:3599 M:252 E: No. of Seq. differs, <211> LENGTH:Input:23 Found:24 SEQ:187
L:3614 M:252 E: No. of Seq. differs, <211> LENGTH:Input:22 Found:23 SEQ:188
L:3629 M:252 E: No. of Seq. differs, <211> LENGTH:Input:21 Found:22 SEQ:189
L:3644 M:252 E: No. of Seq. differs, <211> LENGTH:Input:20 Found:21 SEQ:190
L:3659 M:252 E: No. of Seq. differs, <211> LENGTH:Input:19 Found:20 SEQ:191
L:3673 M:252 E: No. of Seq. differs, <211> LENGTH:Input:18 Found:19 SEQ:192
L:3687 M:252 E: No. of Seq. differs, <211> LENGTH:Input:17 Found:18 SEQ:193
L:3701 M:252 E: No. of Seq. differs, <211> LENGTH:Input:16 Found:17 SEQ:194
L:3714 M:252 E: No. of Seq. differs, <211> LENGTH:Input:15 Found:16 SEQ:195
L:3727 M:252 E: No. of Seq. differs, <211> LENGTH:Input:14 Found:15 SEQ:196
L:3740 M:252 E: No. of Seq. differs, <211> LENGTH:Input:13 Found:14 SEQ:197
L:3753 M:252 E: No. of Seq. differs, <211> LENGTH:Input:20 Found:13 SEQ:198
L:3766 M:252 E: No. of Seq. differs, <211> LENGTH:Input:11 Found:12 SEQ:199
L:3779 M:252 E: No. of Seq. differs, <211> LENGTH:Input:10 Found:11 SEQ:200
L:3792 M:252 E: No. of Seq. differs, <211> LENGTH:Input:9 Found:10 SEQ:201
L:3805 M:252 E: No. of Seq. differs, <211> LENGTH:Input:8 Found:9 SEQ:202
L:3818 M:252 E: No. of Seq. differs, <211> LENGTH:Input:7 Found:8 SEQ:203
L:3831 M:252 E: No. of Seq. differs, <211> LENGTH:Input:6 Found:7 SEQ:204
L:3832 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO:204
L:3884 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO:207
L:3896 M:252 E: No. of Seq. differs, <211> LENGTH:Input:6 Found:5 SEQ:207
L:3897 M:214 E: (33) Seq.# missing, SEQ ID NO:208
L:4233 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
L:4248 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
L:4454 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO:248

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556A

DATE: 04/26/2002

TIME: 11:49:44

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\04262002\I498556A.raw

L:4466 M:252 E: No. of Seq. differs, <211> LENGTH:Input:11 Found:10 SEQ:248
 L:4479 M:252 E: No. of Seq. differs, <211> LENGTH:Input:10 Found:9 SEQ:249
 L:4492 M:252 E: No. of Seq. differs, <211> LENGTH:Input:9 Found:8 SEQ:250
 L:4505 M:252 E: No. of Seq. differs, <211> LENGTH:Input:8 Found:7 SEQ:251
 L:4518 M:252 E: No. of Seq. differs, <211> LENGTH:Input:7 Found:6 SEQ:252
 L:4531 M:252 E: No. of Seq. differs, <211> LENGTH:Input:16 Found:8 SEQ:253
 L:4544 M:252 E: No. of Seq. differs, <211> LENGTH:Input:8 Found:7 SEQ:254
 L:4560 M:252 E: No. of Seq. differs, <211> LENGTH:Input:7 Found:11 SEQ:255
 L:4561 M:214 E: (33) Seq.# missing, SEQ ID NO:256
 L:4575 M:252 E: No. of Seq. differs, <211> LENGTH:Input:5 Found:23 SEQ:257
 L:4717 M:214 E: (33) Seq.# missing, SEQ ID NO:268
 L:4832 M:333 E: Wrong sequence grouping, Amino acids not in groups!
 L:4832 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:5360 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:5920 M:333 E: Wrong sequence grouping, Amino acids not in groups!
 L:5920 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:2
 L:5921 M:333 E: Wrong sequence grouping, Amino acids not in groups!
 L:5921 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:2
 L:5922 M:333 E: Wrong sequence grouping, Amino acids not in groups!
 L:5922 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1

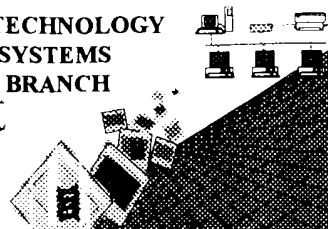
RECEIVED

MAY 9 2002

BIOTECHNOLOGY
SYSTEMS
BRANCH

TECH CENTER 1600 2900

RAW SEQUENCE LISTING
ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/498,556A
Source: OIP
Date Processed by STIC: 4/26/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

RECEIVED

MAY 9 2002

TECH CENTER 1600 2900

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: C9/498,556A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 ✓ Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ✓ No
 ~~Invalid~~ <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.